

# International Workshop on Collaborative Big Data (C-Big 2012)

In conjunction with the 8th IEEE International Conference on Collaborative Computing: Networking, Applications and Worksharing (CollaborateCom 2012)

October 14, 2012, Pittsburgh, Pennsylvania, United States

[Call for Papers](#) | [Important Dates](#) | [Committee](#) | [Tentative Program](#) | [Keynote](#)

## Highlights:

- Dr. Milind Bhandarkar, Chief Architect of Greenplum Labs, EMC will give a talk titled "OpenChorus : Building a Toolchest for Big Data Analytics" as part of the industry session
- C-Big 2012 will feature a best paper award (The paper will be selected by the program committee).
- Prof. Christos Faloutsos from Carnegie Mellon University will give the **keynote talk** at C-Big 2012.
- Selected papers from the workshop will be invited for a special issue in **Elsevier's Information Systems journal**
- To comply with the new deadline of CollaborateCom we extend the paper submission deadline of C-Big 2012 until Aug. 15.

## Call for Papers

The ability to collect, integrate, and analyze data from large number of diverse data sources has increased the amount of data collected and processed by individual organizations on the order of several tens to hundreds of TB. This data can be efficiently utilized for better decision making, improved business intelligence, as well as for enabling new knowledge and services.

Collaborative generation and utilization of large quantities of data, either through crowdsourcing or through efforts of different organizations and groups creates notable research opportunities. The potential benefits of Collaborative Big Data, as well as new and unexpected challenges are still emerging. Managing, processing, and making sense of this data pose new challenges in storage, networking, database management, data mining, knowledge discovery, information security and privacy.

C-Big 2012 brings together researchers and practitioners from around the world to share their experiences on creating, managing, and handling Collaborative Big Data and its benefits.

**The authors of selected papers from the workshop will be invited to submit extended versions of the papers to a special issue in Elsevier's Information Systems journal.**

## Topics of Interest

Topics of interest include (but are not limited to):

- Challenges for collaboration in Big Data utilization
- Network architectures and optimization for Big Data applications
- Collaborative Big Data storage and management in the cloud
- Collaborative Big Data reliability assessment
- Crowdsourcing of the Big Data utilization task
- Large-scale collective intelligence for data integration and data fusion
- Human-centered information fusion and sense-making
- Large-scale process monitoring for handling high data rates
- Big Data for network management
- Big Data and mobile networks including challenges in spectrum management and sensing
- Scalable collaborative graph data processing
- Big Data and social networking including location based social networks
- Security and privacy in Collaborative Big Data
- Applications of Collaborative Big Data

## Important Dates

Paper Submission deadline: ~~August 1, 2012~~ **August 15, 2012**

Notification to authors: ~~September 1, 2012~~ **September 15, 2012**

Camera-ready submission deadline: ~~September 10, 2012~~ **September 25, 2012**

## Paper Submission

Submitted manuscripts should closely reflect the final papers as they will appear in the Proceedings, and should not exceed 8 pages in two-column IEEE proceeding format. We urge the authors to prepare their papers according to the Latex or Microsoft Word templates found at IEEE Author Digital Tool Box, under the "Template for Transactions" section.

All papers are refereed through a single blind process. All papers must be submitted online.

<https://www.easychair.org/conferences/?conf=cbig2012>

All submitted papers will be rigorously reviewed. All accepted papers will be made available in IEEE Xplore and external indexing services (DBLP database, ZB1Math/CompuServe, IO-Port, EI, Scopus, INSPEC, ISI proceeding - pending approval).

## Committees

### General Co-Chairs

Panos K. Chrysanthis, University of Pittsburgh, USA

Ling Liu, Georgia Tech, USA

### Program Co-Chairs

Vladimir Zadorozhny, University of Pittsburgh, USA

Prashant Krishnamurthy, University of Pittsburgh, USA

### Technical Program Committee

Karl Aberer, EPFL, Switzerland

Sujata Banerjee, HP Labs, USA

Ioannis Broustis, AT&T Labs Research, USA

Ugur Cetintemel, Brown University, USA

Keke Chen, Wright State University, USA

Alex Delis, University of Athens, Greece

Minos Garofalakis, Technical University of Crete, Greece

Vanathi Gopalakrishnan, University of Pittsburgh, USA

Christian Jensen, Aarhus University, Denmark

Alexandros Labrinidis, University of Pittsburgh, USA

Harsha V. Madhyastha, University of California, Riverside, USA

Konstantinos Pelechrinis, University of Pittsburgh, USA

Alessandra Sala, Alcatel-Lucent Bell Labs, Ireland

Suresh Singh, Portland State University, USA

Anil Vullikanti, Virginia Tech, USA

Ting Wang, IBM TJ Watson Center, USA

Christo Wilson, Northeastern University, USA

Roberto Zicari, Goethe-Universität Frankfurt, Germany

## Tentative Program

Sunday, October 14, 2012	
8:00 AM – 8:45 AM	Registration
8:50 AM – 9:00 AM	Opening Remarks
9:00 AM – 10:00 AM	Session 1: Keynote Address
	<ul style="list-style-type: none"><li>• <b>Mining Billion-Node Graphs - Patterns and Scalable Algorithms (slides)</b> Prof. Christos Faloutsos</li></ul>
10:00 AM – 10:15 PM	Coffee Break
10:15 AM – 12:15 PM	Session 2: Industry Session
	<ul style="list-style-type: none"><li>• <b>OpenChorus : Building a Toolchest for Big Data Analytics (slides)</b> Milind Bhandarkar, Chief Architect of Greenplum Labs, EMC</li><li>• <b>Real-time analytics using Vertica Analytics Platform (slides)</b> Shilpa Lawande and Stephen Walkauskas, Vertica Group, HP</li></ul>
12:15 PM – 1:30 PM	Lunch
1:30 PM – 3:00 PM	Session 3: Algorithms
	<ul style="list-style-type: none"><li>• <b>Using a Distributed Search Engine to Identify Optimal Product Sets for Use in an Outbreak Detection System</b> Ruhsary Rexit, Fuchiang Tsui, Jeremy Espino, Sahawut Wesaratchakit, Ye Ye and Panos Chrysanthis</li><li>• <b>Towards Efficient Query Processing on Massive Time-Evolving Graphs (slides)</b> Arash Fard, Amir Abdolrashidi, Lakshmesh Ramaswamy and John A. Miller</li><li>• <b>Delta Extraction Optimization for View Maintenance in a Limited Collaborative Environment</b> Zheng Lu, Haijun Liu and Peter Hyland</li></ul>
	Coffee Break
3:30 PM – 5:00 PM	Session 4: Architecture and Applications
	<ul style="list-style-type: none"><li>• <b>Biologically-inspired Network "Memory" for Smarter Networking</b> Bassem Mokhtar and Mohamed Eltoweissy</li><li>• <b>Efficient Processing of Models for Large-scale Shotgun Proteomics Data (slides)</b> Himanshu Grover and Vanathi Gopalakrishnan</li><li>• <b>Robust Decision Engineering: Collaborative Big Data and its Application to International Development/Aid</b> Steve Chan, Wesley Rhodes, Charles Atencio, Brent Ranalli, Anna Miao, Simone Sala, Stephen Serene, Sarah Rumbley, Lisa Sokol, Loren Gary, Caroline Kuo and Marc Clement</li></ul>

## Keynote

Keynote Speaker: **Prof. Christos Faloutsos, Carnegie Mellon University.**

### Mining Billion-Node Graphs - Patterns and Scalable Algorithms

#### Abstract

How do graphs look like? How do they evolve over time? How do rumors and viruses propagate on real graphs? We review some static and temporal 'laws', fast algorithms to spot deviations and outliers, and recent developments on virus propagation and scalable tensor analysis.

#### Biography

Prof. Faloutsos, ACM Fellow, has received the Presidential Young Investigator Award from the National Science Foundation (1989), the Research Contributions Award in ICDM 2006, the SIGKDD Innovations Award (2010), eighteen "best paper" awards (including two "test of time" awards), and four teaching awards. He has served as a member of the executive committee of SIGKDD, he has published over 200 refereed articles, 11 book chapters, and one monograph. He holds six patents and he has given over 30 tutorials and over 10 invited distinguished lectures. His research interests include data mining for graphs and streams, fractals, database performance, and indexing for multimedia and bio-informatics data. He is ranked among the top 50 nurturers in information technology.